



Mitiche, I and Morison, G and Hughes-Narborough, M and Nesbitt, A and Boreham, P and Stewart, B G (2017) Classification of partial discharge signals by combining adaptive local iterative filtering and entropy features. In: 2017 IEEE Conference on Electrical Insulation and Dielectric Phenomena. IEEE Dielectrics and and Electrical Insulation Society, pp. 335-338. (In Press) ,

This version is available at <https://strathprints.strath.ac.uk/62272/>

Strathprints is designed to allow users to access the research output of the University of Strathclyde. Unless otherwise explicitly stated on the manuscript, Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Please check the manuscript for details of any other licences that may have been applied. You may not engage in further distribution of the material for any profitmaking activities or any commercial gain. You may freely distribute both the url (<https://strathprints.strath.ac.uk/>) and the content of this paper for research or private study, educational, or not-for-profit purposes without prior permission or charge.

Any correspondence concerning this service should be sent to the Strathprints administrator: strathprints@strath.ac.uk

The Strathprints institutional repository (<https://strathprints.strath.ac.uk>) is a digital archive of University of Strathclyde research outputs. It has been developed to disseminate open access research outputs, expose data about those outputs, and enable the management and persistent access to Strathclyde's intellectual output.

Classification of Partial Discharge Signals by Combining Adaptive Local Iterative Filtering and Entropy Features

I. Mitiche, G. Morison, M. Hughes-Narborough, A. Nesbitt
Glasgow Caledonian University
70 Cowcaddens Rd
Glasgow, G4 0BA UK

P. Boreham
Doble Engineering
A35, Bere Regis
Wareham, BH20 7JZ UK

B. G. Stewart
Institute of Energy and Environment
University of Strathclyde
204 George St, Glasgow, G1 1XW UK

Abstract- Electro-Magnetic Interference (EMI) is a measurement technique for Partial Discharge (PD) signals which arise in operating electrical machines, generators and other auxiliary equipment due to insulation degradation. Assessment of PD can help to reduce machine downtime and circumvent high replacement and maintenance costs. EMI signals can be complex to analyze due to their nonstationary nature. In this paper, a software condition-monitoring model is presented and a novel feature extraction technique, suitable for nonstationary EMI signals, is developed. This method maps multiple discharge sources signals, including PD, from the time domain to a feature space which aids interpretation of subsequent fault information. Results show excellent performance in classifying the different discharge sources.

I. INTRODUCTION

Partial discharge (PD) diagnosis methods for condition monitoring of insulation materials has received considerable attention recently in research and industry. The most popular method involves the analysis of Phase Resolved PD (PRPD) patterns [1] [2]. However, this method has a few limitations including long data acquisition times and lack of analysis precision. Other previous work in the literature has been successful in the recognition or classification of PD and its different sources with reasonably good performance (see [3] to [9]). However, most of these papers have used simulated or experimental PD data. Moreover, these papers have addressed only the classification of PD sources. In this paper, we aim to investigate the classification of real field Electro Magnetic Interference (EMI) measurements that contain different discharge sources and conditions including PD, corona, arcing, exciter etc. In addition, we develop a novel feature extraction algorithm based on a time-frequency decomposition method, called Adaptive Local Iterative Filtering (ALIF), combined with Permutation Entropy (PE) measure. Multi-Class SVM (MCSVM) is used to classify the different signal conditions for the EMI measurements. The next section describes in more detail the EMI monitoring approach to PD signals used for their measurement. Section III briefly defines

the theory of ALIF and PE algorithms. Section IV defines the mathematical concept of SVM and the application of the feature extraction and classification model to the measured data. Section V describes the measurement procedure. Classification results are presented in Section VI and finally conclusion and future work recommendations are provided in the last section.

II. EMI DESCRIPTION

PD occurs as a localized dielectric breakdown of an electrical insulation system under high voltage stress. The release of energy resulting from PD can take a variety of forms, including EMI. Many types of defects in cables, generators and motors and associated auxiliary equipment generate EMI [10]. The resulting burst of low energy electromagnetic pulses propagates out in all directions from the origin of any discharge event. The EMI signals emitted can be measured and analysed over a broad frequency range. For example, slot discharges can be measured in the range as low as 10's of kHz while loose connections in an isolated phase bus can generate signals up to 100MHz and above. EMI can be radiated and conducted from the discharge site. The part of the energy that is conducted is detected and measured with a split-core high-frequency current transformer (HFCT) over a frequency range of 10 kHz to 100MHz. The EMI measurement of these signals usually follows the spirit of the CISPR-16-1-1 standard to ensure compatibility of measurements and results interpretation across EMI measurement instruments. Convenient locations for the split-core HFCT include, e.g., the neutral connection conduit of a generator stator winding at the grounding transformer, or, the motor cable supply conduit. The measurement of the EMI signals emitted is sensitive to a variety of electrical and mechanical defects. For example, loose or broken stator and rotor bars, slot discharges, winding insulation defects, contamination on the windings, shaft eccentricity, bearing wear, etc. The severity, location and deterioration rate of defects can often be measured long before

detection by other more conventional methods. This is of value to Utilities that desire a system diagnostics approach. EMI analysis can discriminate between different defects and discharge sources and monitors not only activity within, e.g. a generator stator winding but, also, adjacent auxiliary equipment as well. A number of problems have been identified by EMI techniques with exciters, cables, isolated phase bus and transformers, as well as machine stator windings. The resulting EMI spectrum is also unique for each physical location and type of defect within the electrical system [11].

III. FEATURE EXTRACTION

A. Adaptive Local Iterative Filtering (ALIF)

ALIF is a time-frequency signal decomposition which aims to break down the frequency content of a multicomponent signal and produce the time series of each frequency component called the Intrinsic Mode Function (IMF) [12]. This method is similar to Empirical Mode Decomposition (EMD) in that it obtains the IMFs by iteratively subtracting the local average from the signal. However, ALIF employs a filtering technique based on a Partial Differential Equation (PDE) model to calculate the local averages. Moreover, it overcomes EMD issues such as mode mixing [13]. For a signal $y(n) = \text{IMF}_1(n) + \text{IMF}_2(n) + \dots + \text{IMF}_K(n) + R(n)$, which is decomposed into $\text{IMF}_j(n)$; $j=1,2,\dots,K$ plus a residual trend $R(n)$, the ALIF algorithm is briefly summarized as follows.

ALIF ALGORITHM

```

IMF(n) = { }
while the number of extrema  $\geq 2$  do
   $y_1(n) = y(n)$ 
  while the stopping criteria is not satisfied do
    Calculate filter length  $l_j(n)$  and filter
    coefficients for  $y(n)$ 
    
$$y_{j+1}(n) = y_j(n) - \sum_{m=-l_j(n)}^{l_j(n)} h_j(n, m) \cdot y_j(n + m)$$

     $j = j + 1$ 
  end while
   $\text{IMF}(n) = \text{IMF}(n) \cup y_j(n)$ 
   $y(n) = y(n) - y_j(n)$ 
end while
 $\text{IMF}(n) = \text{IMF}(n) \cup y(n)$ 

```

where $h_j(n, m) \in \mathbb{R}$, $n \in [-l_j(n), l_j(n)]$ are the coefficients of a low pass filter at point n with length $2l_j(n) + 1$.

B. Permutation Entropy (PE)

PE is a measure of complexity in a time series with nonlinear characteristics [14] and is based on Shannon

Entropy which estimates the meaningful information in data[15]. PE is suitable for nonstationary and nonlinear time series, it is also robust against low frequency artifacts, fast to compute and stable for measuring randomness in time series data[16]. These advantages overcome the limitations of the basic entropy based measures. PE theory is described as follows.

Based on a time series $\{x_i\}$; $i=1,2,\dots,N$, we construct vectors of m points plus $m!$ possible ordinal patterns $\pi_i^{m,\tau}$, where τ is a delay. Next, we seek the occurrence of these patterns in the time series and hence calculate the permutation probability for each pattern as:

$$\pi_i^{m,\tau} = \frac{\#\{i \mid 0 \leq i \leq N - m, (x_{i+1}, \dots, x_{i+m}) \text{ has type } \pi_i^{m,\tau}\}}{N - m + 1}. \quad (1)$$

Finally PE value is then calculated using Shannon's Entropy formula:

$$H(n) = - \sum_{i=1}^{k \leq m!} p(\pi_i^{m,\tau}) \cdot \log(p(\pi_i^{m,\tau})). \quad (2)$$

IV. SUPPORT VECTOR MACHINE (SVM)

SVM is used in this work as a binary classification method. It separates two different classes with an optimum line called a hyperplane[17]. The main advantages of SVM include ability to handle large data features, and high classification accuracy compared to other classification techniques such as Neural Networks (NN) and random forests [18].

The theory of SVM is described as follows. The data set x_i and its labels y_i ; $i=1,2,\dots,L$ with L samples, is assumed to contain points of two classes "1" and "-1". Each point is placed in a feature space that is separated by the hyperplane defined as:

$$f(x) = w \cdot x + b = 0, \quad (3)$$

where w is an L -dimensional vector and b is a scalar, which both play an important role in defining the hyperplane's position. The hyperplane will pass by the origin if $b = 0$ otherwise the margin is increased. Two parallel hyperplanes are created for each class and are defined in (8) and (9) for the first and second class respectively.

$$w \cdot x + b = 1. \quad (4)$$

$$w \cdot x + b = -1. \quad (5)$$

The distance between the hyperplanes is $2/|w|$. This margin width can be maximized by minimizing $|w|$ which results in the criteria for the first and second class respectively as follows.

$$w \cdot x_i + b \geq 1 \quad ; \quad w \cdot x_i + b \geq -1.$$

This will ensure that the points from each class will not exceed their respective hyperplane. The points that lie on the hyperplane position are called support vectors. The hyperplane is calculated by solving the optimization problem defined in (6) with taking into consideration an error penalty C which optimizes the margin width and classification error in training the SVM,

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^M \zeta_i, \quad (6)$$

$$\text{s.t.} \begin{cases} y_i (w^T \cdot x_i + b) \geq 1 - \zeta_i \\ \zeta_i \geq 0, \quad i = 1, \dots, M \end{cases},$$

where ζ_i represents the distance between the margin and the samples that overstep the margin. Equation (6) is solved through the Lagrangian approach [19]. The solution is expressed in (7) where α_i is a parameter that expresses w in solving (6).

$$f(x) = \text{sign} \left(\sum_{i,j=1}^M \alpha_i y_i (x_i x_j) + b \right). \quad (7)$$

The challenges that may be faced during the SVM training process are data over-fitting and computational errors. The first challenge can be addressed by employing a kernel function which calculates the dot product of the feature space as in (8) and which is described in more detail in [17].

$$K(x_i x_j) = (\Phi^T(x_i) \cdot \Phi(x_j)), \quad (8)$$

where $\Phi(x) = (\phi_1(x), \dots, \phi_l(x))$ is a nonlinear vector function and l is the feature space dimension. Equation (7) can then be reformulated as:

$$f(x) = \text{sign} \left(\sum_{i,j=1}^M \alpha_i y_i (\Phi^T(x_i) \cdot \Phi(x_j)) + b \right). \quad (9)$$

In this paper, SVM is used as part of the MCSVM algorithm in order to classify more than two discharge sources. This is achieved by using the One-Against-One (OAO) approach which builds $k(k-1)/2$ models each one is trained on two classes, “1” and “-1”, using SVM binary classification. The testing is performed by a “Max Win” voting strategy, in that if the test sample is close to one of the classes then the vote for that class is incremented by one. The sample is classified in the category with the highest vote [20].

As illustrated in Fig 1., the approach is to first decompose the signal using ALIF then extract PE feature from each IMF, the feature vector for each signal is then implemented in MCSVM. An example of PN signal with its decomposed

IMFs and relative PE values is illustrated in Fig. 2. The training/testing was achieved using a ten-fold hold-on cross validation technique where the MCSVM is trained on 90% of the data and is tested on the remaining 10% in ten iterations; each uses a different batch of data. The total classification accuracy is the average accuracy of all iterations.

V. DATA MEASUREMENTS

The data was collected during the operation of the assets by placing the HFCT around neutral earth cable in order to pick up the electromagnetic discharges. The data signals were then recorded at a sampling rate of 24 kHz through a PDS200 measurement device which follows the CISPR16 standard, using EMC type filtering. The PDS200 is used for PD surveying and has the ability to detect and analyze Radio Frequency Interference (RFI) as well as EMI radiation. It also provides the frequency spectrum of the recorded signals using the EMI standard quasi peak detection method. “EMI experts” select the frequency peaks that are suspicious of activity and which may represent a fault, analyze and identify the condition or fault and label the time resolved signals that correspond to these frequency peaks. These labels help in training the classification algorithm. The signals were collected from three different power plants which are described as follows.

- 1) Site 1: The data was measured at the neutral earth cable of a 661MVA hydrogen/water cooled synchronous generator. A total of 13 signals were selected with 10s duration and contain 500 frequency cycles each.
- 2) Site 2: similar to the previous site, the measurements were taken at the neutral earth of different assets where 6 signals were selected from a General Step-Up (GSU) transformer, 2 signals from an Isolated Phase Bus (IPB) and one signal from a steam turbine transformer. All files also have a capture duration of 10s and 500 cycles.
- 3) Site 3: The data was measured at the neutral earth cable of the generator from which 7 signals were selected and 1 signal was selected from a Steam Turbine Generator (STG). All files are similar in capture to the above.

The discharge sources identified in each site are presented in Table I. The faults selected as common between sites are underlined in the table. As a consequence, the data selected for “common between sites” is a subset of the data collected across the three sites. The signals containing each condition originate from different assets including the generator, IPB and GSU.

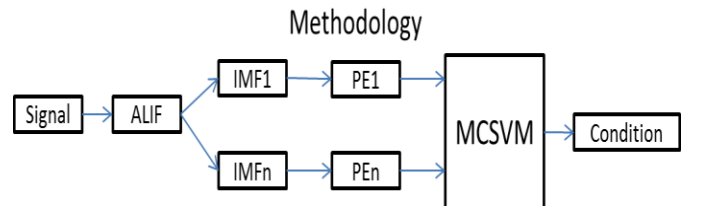


Fig. 1 Classification approach

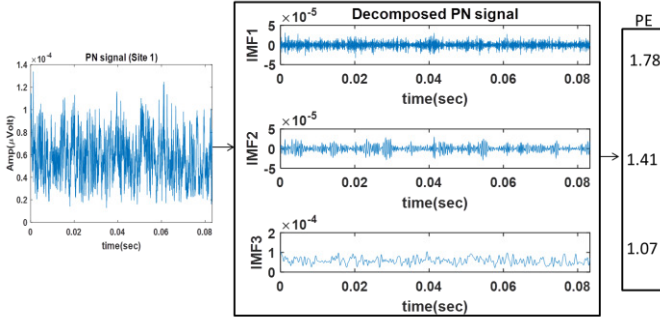


Fig. 2 Application example of the approach

TABLE I
DETECTED FAULTS

Site	Fault
1	Corona, noise, Process Noise (PN), minor PD (mPD)
2	mPD+minor Arcing (mA), PD+ mA, PD+arcing, <u>PD</u> , <u>PN</u> , data modulation(DM)
3	Exciter(E)+mPD+ DM, PN, <u>PD</u> , E+mPD, E, NonVariable Frequency Drive

VI. RESULTS

Classification accuracy for each scenario is illustrated in Table II. It is shown that high classification accuracy is obtained. The highest accuracy (96%) was achieved in the “common between sites” scenario. This shows that the different discharge sources, including PD and combined conditions such as PD+mA, can be distinguished regardless of the site in which the data was collected, whether within the same site or across multiple sites. This means that the conditions exhibit unique patterns. The other significant output is the successful classification of the signals measured on different assets. For instance, PD was collected at GSU and IPB units in site 2 and it was also measured from generators in site 1 and 3. These results appear to be consistent with other existing PD classification work that uses a similar approach.

VI. CONCLUSION

This work introduces a classification technique to PD EMI signals for the first time. Initially, the relevant fault information was extracted using ALIF and PE features. Then, MCSVM was used to classify different discharge sources collected in three power plant sites. Results bring to conclusion that it is possible to establish discharge sources classification in real field EMI captured data using this novel proposed approach. The other significant contribution is the successful classification of common sources between different sites. As future work, this study may be applied for second stage classification to identify PD fault locations.

TABLE II
CLASSIFICATION RESULTS

Site	Classification accuracy
1	83%
2	86%
3	92%
Common between sites	96%

REFERENCES

- [1] F. Álvarez, F. Garnacho, A. Khamlichi, and J. Ortego, ‘Classification of partial discharge sources by the characterization of the pulses waveform’, in 2016 IEEE International Conference on Dielectrics (ICD), 2016, vol. 1, pp. 514–519.
- [2] R. Altenburger, C. Heitz, and J. Timmer, ‘Analysis of Phase-Resolved Partial Discharge Patterns of Voids Based on a Stochastic Process Approach’, J. Phys. D: Appl. Phys., no. 35, pp. 1149–1163, 2002.
- [3] L. V. Badiou, W. Koltunowicz, U. Broniecki, and B. Battle, ‘Increased Operation Reliability Through PD Monitoring of Stator Winding’, in 13th INSUCON Conference, Birmingham, UK, 2017.
- [4] W. J. K. Raymond, H. A. Illias, and A. H. Abu Bakar, ‘Classification of Partial Discharge Measured under Different Levels of Noise Contamination’, PLoS ONE, vol. 12, no. 1, Jan. 2017.
- [5] V. Chatpattananan, N. Pattanadech, and K. Vicetjindavat, ‘PCA-LDA for Partial Discharge Classification on High Voltage Equipment’, in 2006 IEEE 8th International Conference on Properties applications of Dielectric Materials, 2006, pp. 479–481.
- [6] N. Pattanadech, P. Nimsanong, S. Potevijkul, P. Yuthagowith, and S. Polmai, ‘Partial discharge classification using probabilistic neural network model’, in 2015 18th International Conference on Electrical Machines and Systems (ICEMS), 2015, pp. 1176–1180.
- [7] E. M. Lalitha and L. Satish, ‘Wavelet analysis for classification of multi-source PD patterns’, IEEE Trans. Dielectr. Electr. Insul., vol. 7, no. 1, pp. 40–47, Feb. 2000.
- [8] K. Wang, J. Li, S. Zhang, Y. Qiu, and R. Liao, ‘Time-frequency features extraction and classification of partial discharge UHF signals’, in 2014 International Conference on Information Science, Electronics and Electrical Engineering, 2014, vol. 2, pp. 1231–1235.
- [9] J. A. Hunter, L. Hao, P. L. Lewin, D. Evagorou, A. Kyprianou, and G. E. Georgiou, ‘Comparison of two partial discharge classification methods’, in Electrical Insulation (ISEI), Conference Record of the 2010 IEEE International Symposium on, 2010, pp. 1–5.
- [10] J. E. Timperley and J. M. Vallejo, ‘Condition assessment of electrical apparatus with EMI diagnostics’, in 2015 IEEE Petroleum and Chemical Industry Committee Conference (PCIC), 2015, pp. 1–8.
- [11] J. E. Timperley, J. M. Vallejo, and A. Nesbitt, ‘Trending of EMI data over years and overnight’, in 2014 IEEE Electrical Insulation Conference (EIC), 2014, pp. 176–179.
- [12] A. Cicone, J. Liu, and H. Zhou, ‘Adaptive local iterative filtering for signal decomposition and instantaneous frequency analysis’, Appl. Comput. Harmon. Anal., vol. 41, no. 2, pp. 384–411, 2016.
- [13] N. E. Huang et al., ‘The Empirical Mode Decomposition and The Hilbert Spectrum for Nonlinear and Non-Stationary Time Series Analysis’, Proc. R. Soc. Lond. Math. Phys. Eng. Sci., vol. 454, no. 1971, pp. 903–995, 1998.
- [14] C. Bandt and B. Pompe, ‘Permutation Entropy: A Natural Complexity Measure for Time Series’, Phys. Rev. Lett., vol. 88, no. 17, p. 174102, Apr. 2002.
- [15] S. D. S. P. Rathie, ‘Shannon, Levy, and Tsallis: a note’, Appl. Math. Sci., vol. 2, no. 28, pp. 1359–1363, 2008.
- [16] A. M. M. Riedl and N. Wessel, ‘Practical Considerations of Permutation Entropy’, Eur. Phys. J. Spec. Top., vol. 222, no. 2, pp. 249–262, 2013.
- [17] V. Vapnik, The Nature of Statistical Learning Theory. Springer, 1995.
- [18] S. G. Lesniak JM. Hupse R. ..Blanc R. ..Karssemeijer N., ‘Comparative Evaluation of Support Vector Machine Classification for Computer Aided Detection of Breast Masses in Mammography’, Phys. Med. Biol., vol. 57, pp. 2560–2574, Aug. 2012.
- [19] A. Widodo and B.-S. Yang, ‘Support Vector Machine in Machine Condition Monitoring and Fault Diagnosis’, Mech. Syst. Signal Process., vol. 21, no. 6, pp. 2560–2574, 2007.
- [20] C.-W. Hsu and C.-J. Lin, ‘A Comparison of Methods for Multi-class Support Vector Machines’, IEEE Trans. Neural Netw., vol. 13, no. 2, pp. 415–425, 2002.